Avocet Consulting

TECHNICAL MEMORANDUM

October 5, 2004

TO: Mike Poulsen and Jennifer Peterson, DEQ
Brett Betts, Ecology
Taku Fuji, Kennedy Jenks

FROM: Teresa Michelsen, Avocet Consulting

RE: Additional reliability analysis for WA Freshwater SQGs

This memorandum provides the results of additional reliability analysis of the candidate SQG sets for freshwater sediments in WA and OR, as requested by DEQ. Specifically, DEQ requested that the following suite of reliability parameters be added at the Stat-Only, SQS, and CSL levels:

- **Predicted Hit Sensitivity** = correctly predicted hits / total predicted hits

- **False Predicted Hits** = incorrectly predicted hits / total predicted hits
       Predicted Hit Sensitivity = 1 – False Predicted Hits

- **Predicted No-Hit Efficiency** = correctly predicted no-hits / total predicted no-hits

- **False Predicted No-Hits** = incorrectly predicted no-hits / total predicted no-hits
       Predicted No-Hit Efficiency = 1 – False Predicted No-Hits

A figure provided by DEQ illustrating these reliability measures is attached as it is very useful in visualizing the various reliability parameters. Note that Predicted Hit Sensitivity is the same as the 1988 Efficiency previously calculated, and the False Predicted Hits and False Predicted No-Hits are derived from two of the other parameters. Therefore, the only new reliability calculations were for the Predicted No-Hit Efficiency, the parameter which DEQ is particularly interested in utilizing for making decisions about SQGs.

DEQ requested that Tables 3-8, 3-9, and 3-10 from the Phase II report be updated with these additional parameters, and provided an example table based on the assumption that the number of hits and no-hits were equal. As indicated by DEQ, the additional calculations were conducted using the unsummed data set and a comparison to control for the three levels of effects described above. The revised tables are attached.

The following observations can be made:

- As was the case previously, the Stat-Only and SQS levels are very similar. This is to be expected, since the SQS level of effects was based on the minimum detectable difference expected in the bioassays used. The Stat-Only level has slightly poorer performance in the Predicted No-Hit Efficiency, therefore, I would recommend that the SQS level be used in preference to the Stat-Only level.

- The Predicted Hit Sensitivity and the Predicted No-Hit Efficiency are not the same as those in the example table provided by DEQ, because the percentage of hits and no-hits is not equal. Deviations from DEQ's projections are substantial at the SQS level, but much smaller at the CSL level. In general, at the SQS level, the Predicted Hit Sensitivity is higher and the Predicted No-Hit Efficiency is lower than projected.

- The Predicted No-Hit Efficiency improves substantially for the CSL level over the SQS level. Thus, if this criterion is most important to you, you may wish to choose a more sensitive SQG set for the SQS level, and a less sensitive SQG set for the CSL level. This would also seem to be justified by the conceptual goals for these levels.

Table 3-8. Floating Percentile Results for Statistical Significance Only Effects Level
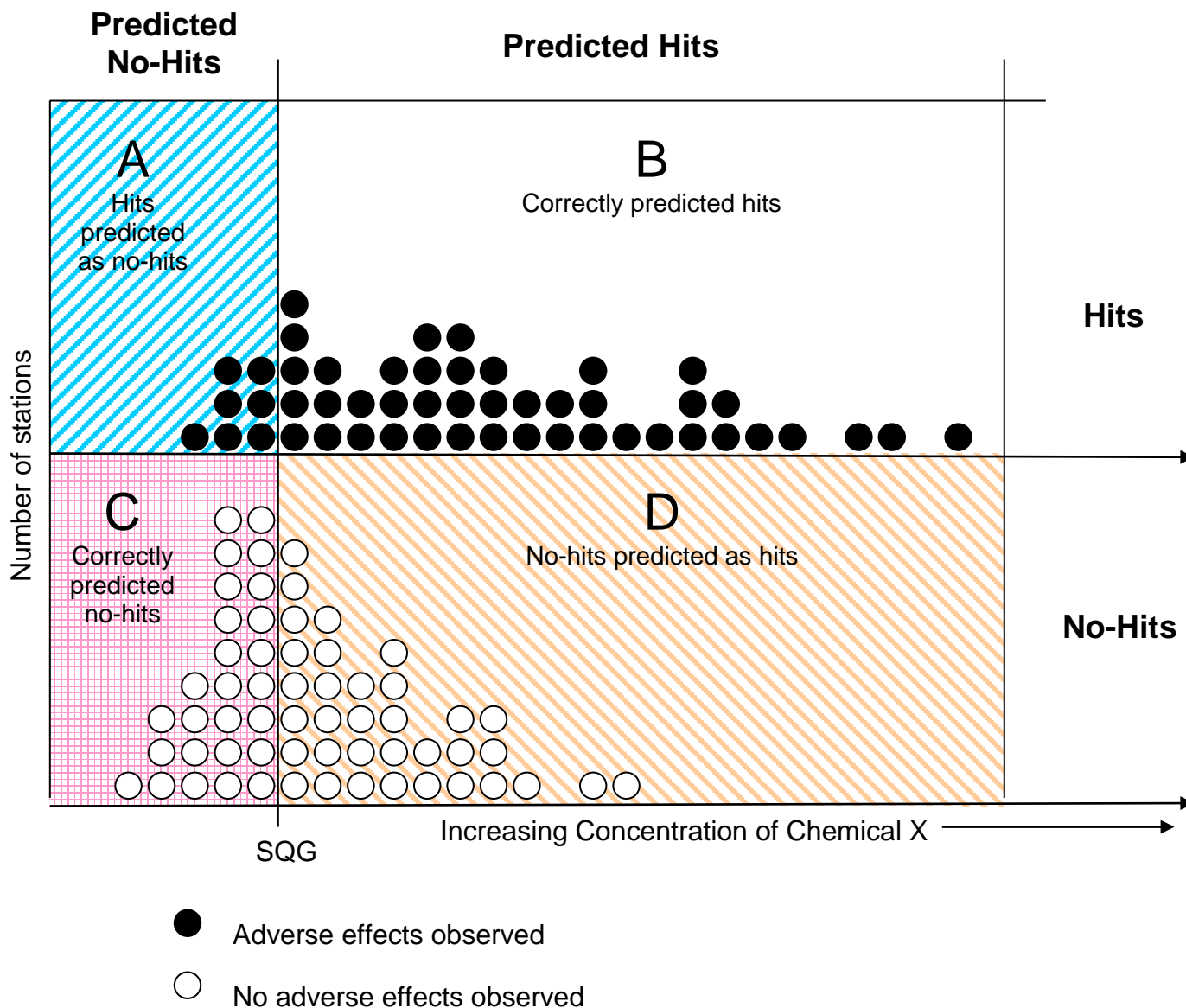
Measure of Reliability (%)	95% Sensitivity	90% Sensitivity	85% Sensitivity	80% Sensitivity	75% Sensitivity
Sensitivity	96	90	85	80	75
False Negatives	4	10	15	20	25
Efficiency	45	61	74	80	84
False Positives	55	39	26	20	16
Predicted Hit Sensitivity	82	86	89	92	93
False Predicted Hits	18	14	11	8	7
Predicted No-Hit Efficiency	79	69	65	60	56
False Predicted No-Hits	21	31	35	40	44
Overall Reliability	82	82	82	81	78

Table 3-9. Floating Percentile Results for the SQS Effects Level

Measure of Reliability (%)	95% Sensitivity	90% Sensitivity	85% Sensitivity	80% Sensitivity	75% Sensitivity
Sensitivity	95	90	85	80	75
False Negatives	5	10	15	20	25
Efficiency	43	56	74	80	85
False Positives	57	44	26	20	15
Predicted Hit Sensitivity	79	82	88	90	92
False Predicted Hits	21	18	12	10	8
Predicted No-Hit Efficiency	80	71	69	63	60
False Predicted No-Hits	20	29	31	37	40
Overall Reliability	80	80	82	80	78

Table 3-10. Floating Percentile Results for the CSL Effects Level

Measure of Reliability (%)	95% Sensitivity	90% Sensitivity	85% Sensitivity	80% Sensitivity	75% Sensitivity
Sensitivity	95	90	85	80	75
False Negatives	5	10	15	20	25
Efficiency	50	63	74	77	79
False Positives	50	37	26	23	21
Predicted Hit Sensitivity	66	71	76	77	78
False Predicted Hits	34	29	24	23	22
Predicted No-Hit Efficiency	92	86	83	79	76
False Predicted No-Hits	8	14	17	21	24
Overall Reliability	73	77	80	79	78

Predicted No-Hits

Predicted Hits

A
Hits predicted as no-hits

B
Correctly predicted hits

Hits

Number of stations

C
Correctly predicted no-hits

D
No-hits predicted as hits

No-Hits

SQG

Increasing Concentration of Chemical X ⟶

● Adverse effects observed

○ No adverse effects observed

Sensitivity = B / (A + B)
False Negatives = A / (A + B)

Efficiency = C / (C + D)
False Positives = D / (C + D)

Predicted-Hit Sensitivity = B / (B + D)
False Predicted Hits = D / (B + D)

Predicted-No-Hit Efficiency = C / (A + C)
False Predicted No-Hits = A / (A + C)

Figure 1. Accuracy Indicator Definitions